FBS31 – GlobalFiler™ Interpretation Guidelines

Table of Contents

1. Scope

1.1. This procedure outlines the guidelines used to interpret the results obtained from the GeneMapper® ID-X (GMID-X) analysis and STRmix™ software programs for samples amplified with GlobalFiler™. These guidelines ensure that conclusions in casework reports are scientifically supported by the analytical data (including data from appropriate standards and controls) and interpretations are made objectively, are consistent from analyst to analyst, and are within previously validated limits.

2. Background

2.1. These practices set forth the FSL FBU's approach for conducting and documenting the interpretation of DNA profiles generated from evidence in conformance with the requirements of the Department of Forensic Sciences (DFS) *Forensic Science Laboratory (FSL) Quality Assurance Manual*, the accreditation standards under ISO/IEC 17025 (current revision), and the supplemental standards set by the FSL's accrediting body, as well as the guidance provided in the *Scientific Working Group on DNA Analysis Methods*

(SWGDAM) Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Laboratories and the Federal Bureau of Investigation *Quality Assurance Standards (QAS) for Forensic DNA Testing Laboratories.*

2.2. The interpretation of results in casework is a matter of professional judgment and expertise. Not every situation can, or should, be covered by a pre-set rule; nor is it expected that competent analysts will always be in full agreement in all instances. However, in an effort to achieve uniformity in approach, the Laboratory has developed and adheres to these basic guidelines for the interpretation of analytical results. The principles are based on the manufacturer's user guides and manuals, internal validation studies, scientific literature references, casework experiences, SWGDAM documents, the FBI's QAS, guidance from the DFS Science Advisory Board, and input from the forensic community at large.

3. Safety

3.1. Not applicable

4. Materials Required

4.1. GeneMapper® ID-X Software, version 1.6

4.2. STRmix™ Software, version 2.4

4.3. Windows-based computer capable of running the software

5. Standards and Controls

5.1. GlobalFiler™ Loci

Locus designation	Chromosome location	Alleles included in Allelic Ladder	Dye label	DNA Control 007
D3S1358	3p21.31	9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20	6-FAM™	15, 16
vWA	12p13.31	11,12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24		14, 16
D16S539	16q24.1	5, 8, 9, 10, 11, 12,13, 14, 15		9, 10
CSF1PO	5q33.3-34	6, 7, 8, 9, 10, 11, 12, 13, 14, 15		11, 12
TPOX	2p23-2per	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15		8, 8
Y indel	Yq11.221	1, 2	VIC™	2
Amelogenin	X: p22.1-22.3 Y: p11.2	X, Y		X, Y
D8S1179	8q24.13	5, 6, 7, 8, 9 10, 11, 12, 13, 14, 15, 16, 17, 18, 19		12, 13
D21S11	21q11.2-q21	24, 24.2, 25, 26, 27, 28, 28.2, 29, 29.2, 30, 30.2, 31, 31.2, 32, 32.2, 33, 33.2, 34, 34.2, 35, 35.2, 36, 37, 38		28, 31
D18S51	18q21.33	7, 9, 10, 10.2, 11, 12, 13, 13.2, 14, 14.2, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27		12, 15
DYS391	Yq11.21	7, 8, 9, 10, 11, 12, 13		11
D2S441	2p14	8, 9, 10, 11, 11.3, 12, 13, 14, 15, 16, 17	NED™	14, 15
D19S433	19q12	6, 7, 8, 9, 10, 11, 12, 12.2, 13, 13.2, 14, 14.2, 15, 15.2, 16, 16.2, 17, 17.2, 18.2, 19.2		14, 15
TH01	11p15.5	4, 5, 6, 7, 8, 9, 9.3, 10, 11, 13.3		7, 9.3
FGA	4q28	13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 26.2, 27, 28, 29, 30, 30.2, 31.2, 32.2, 33.2, 42.2, 43.2, 44.2, 45.2, 46.2, 47.2, 48.2, 50.2, 51.2		24, 26

Locus designation	Chromosome location	Alleles included in Allelic Ladder	Dye label	DNA Control 007
D22S1045	22q12.3	8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	TAZ™	11, 16
D5S818	5q21-31	7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18		11, 11
D13S317	13q22-31	5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16		11, 11
D7S820	7q11.21-22	6, 7, 8, 9, 10, 11, 12, 13, 14, 15		7, 12
SE33	6q14	4.2, 6.3, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 20.2, 21, 21.2, 22.2, 23.2, 24.2, 25.2, 26.2, 27.2, 28.2, 29.2, 30.2, 31.2, 32.2, 33.2, 34.2, 35, 35.2, 36, 37		17, 25.2
D10S1248	10q26.3	8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19	SID™	12, 15
D1S1656	1q42.2	9, 10, 11, 12, 13, 14, 14.3, 15, 15.3, 16, 16.3, 17, 17.3, 18.3, 19.3, 20.3		13, 16
D12S391	12p13.2	14, 15, 16, 17, 18, 19, 19.3, 20, 21, 22, 23, 24, 25, 26, 27		18, 19
D2S1338	2q35	11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28		20, 23

5.2. GeneScan™-600 LIZ® Size Standard – contains DNA fragments of known sizes (20, 40, 60, 80, 100, 114, 120, 140, 160, 180, 200, 214, 220, 240, 250, 260, 280, 300, 314, 320, 340, 360, 380, 400, 414, 420, 440, 460, 480, 500, 514, 520, 540, 560, 580 and 600) that are used in-lane to estimate sizes (base pair - bp) of STR products.

5.3. Positive Amplification Control (DNA Control 007) — the positive amplification control is a human male DNA sample with a known autosomal STR genotype (see Section 5.1 above) included in the GlobalFiler™ PCR Amplification Kit. The positive amplification control ensures that the amplification process is working properly. See above diagram for allele calls.

5.4. Negative Amplification Control – the negative amplification control is a check for contamination during set up of the PCR amplification reaction. It essentially monitors the "environment" in that process for possible sources of contamination.

5.5. Reagent Blank (extraction control) – the reagent blank is a check for possible contamination of the extraction reagents by human DNA or by amplified STR product.

6. Procedures

6.1. Overall Interpretation Outline

6.1.1. The analytical controls are evaluated and confirmed. (refer to FBS39)

6.1.2. Artifacts inherent to the PCR process as well as those generated by the detection instrumentation are isolated and eliminated. (refer to FBS39)

6.1.3. The overall profile is evaluated to assess its quality.

6.1.4. The overall profile is defined as a single source or mixture.

6.1.5. The potential number of contributors is determined.

6.1.6. Profiles may be deconvoluted and/or compared to the known profiles submitted in connection with the case in order to make a forensic conclusion regarding the potential exclusion or inclusion of each individual as a contributor(s) of/to each sample.

6.2. Profile Assessment

6.2.1. **Peak Evaluation** – as a profile is evaluated, each peak identified by the GMID-X software must be checked to confirm it is a valid allele call. Typical characteristics and requirements related to the shape, size and location of a true allele are listed below along with a variety of recognizable artifacts.

6.2.1.1. Alleles – a reportable, true allele peak is defined as a distinct, triangular section of an electropherogram that is equal to or higher than the analytical threshold (AT). The analytical threshold is the minimum signal at which a peak can reliably be distinguished from noise. Based on internal validation studies, an AT of 90 rfu has been established for samples amplified with GlobalFiler™ (injected for 24 seconds at 1.2kV on the 3500xL instrument or 15 seconds at 1.2kV on the 3500 instrument).

6.2.1.1.1. Based on an initial review of results, when appropriate, and the quantity of sample allows, the analyst may opt to concentrate the extract and re-amplify, re-amplify with more and/or less template DNA, and/or re-extract. When a sample is analyzed more than one time, the first page of the electropherogram used for reporting purposes will be clearly marked. If replicate amplifications will be reported for a sample, the first page of both electropherograms will be marked.

6.2.1.2. Off Ladder Alleles (OL): An off ladder allele occurs when an allele, in relationship to the allelic ladder, is between two allelic ladder peaks or before or after the peaks in the allelic ladder and not thought to be due to migration effects. In this instance, the peak is designated as “OL” by the GMID-X software.

6.2.1.2.1. When an OL allele occurs between two allelic ladder peaks, this allele will be considered to be a microvariant. The analyst will assign an allele designation based on the base pair value of the allele. The allele will be assigned a designation of the lower complete repeat value followed by the number of bases in the incomplete repeat (e.g., an allele which migrates two bases above the D7S820 11 allele will be designated as D7S820 11.2).

6.2.1.2.2. For single source samples, when an OL allele is seen between two loci and either the locus to the left or the right contains two peaks, the allele will be considered to belong to the locus containing the single peak. Based on the base pair value of the allele, when possible, an allele designation will be assigned (e.g., an allele is noted between D8S1179 and D21S11 and a two allele pattern is noted only at the D21S11 locus, the allele is assigned to the D8S1179 locus. Further, given that the base pair value for the allele in this example is approximately 4 base pairs longer than the longest allele in the D8S1179 ladder (allele 19), the allele in question may be designated as 20). The allele may also be designated as > or < the closest allele in the allelic ladder on the electropherogram print-out. If an allele is detected between FGA 33.2 and 42.2, the allele will be designated as >33.2. When an allele is seen above the largest or below the smallest allele of an allelic ladder, the allele will be designated as either > or < the respective ladder allele (e.g., an allele which migrates below the D3S1358 12 allele will be designated as D3S1358 <12) on the electropherogram print-out. For STRmix™ purposes, the locus will be ignored.

6.2.1.2.3. The sample will be re-injected or re-run to confirm the allele designation of the microvariant. The additional view of the re-injected/re-run sample electropherogram will be designated as listed in section 6.2.1.2.1 and printed/maintained in the case file. If the entire sample will be deemed uninterpretable, confirmation is not required.

NOTE: If other samples from the same case exhibit the same microvariant, and the samples have been run on different 3500/3500xL injections, this will satisfy the requirement for an additional view.

6.2.1.2.4. For single source samples, if an OL allele is seen between two loci and the surrounding loci each contain only a single peak (or two peaks), both loci will be deemed inconclusive.

6.2.1.2.5. When an off ladder allele is seen between two loci in a mixture profile, both loci will typically be deemed inconclusive unless the allele can be assigned to a

locus based on peak heights or an assumed contributor profile. Consult with the Technical Leader as needed.

6.2.1.3. Artifacts/Non-Allelic Peaks:

6.2.1.3.1. Stutter Products – a stutter peak is a well-characterized PCR artifact that refers to the appearance of a minor peak one repeat unit smaller (or less frequently, one repeat larger) than the major STR product. All of the loci in the GlobalFiler™ kit are tetranucleotide repeat units, except for the trinucleotide repeat locus D22S1045, and the insertion/deletion locus Y Indel. Thus when present, the reverse stutter peak will be one repeat shorter (or N-1), while the forward stutter peak will be one repeat longer (or N+1). Two loci (SE33 and D1S1656) also exhibit half-repeat reverse stutter (N-0.5). In general for each locus, the percent reverse stutter increases with allele length.

6.2.1.3.1.1. The proportion of the stutter product relative to the main allele (percent stutter) is measured by dividing the height of the peak in the stutter position by the height of the main allele peak.

6.2.1.3.1.2. The following table lists the reverse stutter (N-1) percentages expected for each locus as determined by internal validation studies for the GlobalFiler™ Amplification Kit.

Locus	N-1 Stutter %	Locus	N-1 Stutter %	Locus	N-1 Stutter %	Locus	N-1 Stutter %	Locus	N-1 Stutter %
D3S1358	11.4349	Y indel	-	D2S441	8.2702	D22S1045	15.1961	D10S1248	11.2523
vWA	10.9288	Amelogenin	-	D19S433	9.9798	D5S818	9.1103	D1S1656	12.1085
D16S539	9.6511	D8S1179	9.2054	TH01	3.8354	D13S317	10.1049	D12S391	14.9596
CSF1PO	9.2886	D21S11	10.0098	FGA	11.4278	D7S820	8.7321	D2S1338	12.6825
TPOX	4.9691	D18S51	12.6458			SE33	14.8328		
		DYS391	7.886						

6.2.1.3.1.3. The following table lists the forward stutter (N+1) percentages expected for each locus as determined by the internal validation studies for the GlobalFiler™ Amplification Kit:

Locus	N+1 Stutter %	Locus	N+1 Stutter %	Locus	N+1 Stutter %	Locus	N+1 Stutter %	Locus	N+1 Stutter %
D3S1358	1.8409	Y indel	-	D2S441	1.4848	D22S1045	7.6419	D10S1248	2.1
vWA	1.7052	Amelogenin	-	D19S433	1.5778	D5S818	1.7784	D1S1656	1.7712
D16S539	1.6176	D8S1179	1.9704	TH01	0.9151	D13S317	1.5304	D12S391	3.2633
CSF1PO	2.1371	D21S11	2.0882	FGA	1.9467	D7S820	1.0587	D2S1338	2.8583
TPOX	0.7352	D18S51	3.2504			SE33	2.7437		
		DYS391	-						

6.2.1.3.1.4. The following table lists the half-repeat reverse stutter (N-0.5) percentages expected for two loci as determined by the internal validation studies for the GlobalFiler™ Amplification Kit:

Locus	N-0.5 Stutter %
SE33	4.2
D1S1656	2.4278

6.2.1.3.1.5. The stutter cut-off values for each locus determined during the internal validation studies are included in the filtering step of the initial GMID-X analysis. A peak in a stutter position that is above the stutter cut-off value will not be filtered. Minor peaks in a stutter position that have not been filtered will remain labeled and should be further evaluated. These peaks may be called elevated stutter at the analyst's discretion based on the specific circumstances.

6.2.1.3.1.6. Stutter peaks are often elevated in profiles with higher rfu values. In particular, the percent stutter for peaks that are off-scale may be unusually high.

6.2.1.3.1.7. Likewise, samples with stochastic effects due to very low input DNA may also have stutter peaks that are elevated.

6.2.1.3.1.8. If there is no indication of a mixture in a profile, the labeled minor peaks in N-1, N-0.5 and N+1 stutter positions can

most often be attributed to elevated stutter. If a mixture is observed, labeled minor peaks in N-1, N-0.5 and N+1 stutter positions must be carefully evaluated. When they cannot be attributed to elevated stutter, they will be taken into consideration during the mixture interpretation process.

6.2.1.3.1.9. When a peak is present between a parent peak and its possible stutter peak, the stutter filter will not be applied by the software. Accordingly, any such peaks in stutter position must be assessed to determine if they may be unfiltered stutter (e.g., when the result at D21S11 is 33.2, 34, 34.2, the 33.2 peak must be assessed to see if it meets the criteria for stutter corresponding to the 34.2 peak, or when minus A is present, a peak in the stutter position must be assessed manually).

6.2.1.3.2. Spikes – a spike is a random non-reproducible artifact usually observed in one or more dye colors at the same base pair location, generally seen as a tall thin peak that does not display proper peak morphology. Peak height (rfu) usually varies between dye colors. Samples and controls containing spikes within the defined analysis range (60 – 460bp) do not need to be re-injected or re-run if the spike does not interfere with the sizing of the sample.

6.2.1.3.3. Dye Blobs –Dye blobs usually appear as a broad peak in a single color. Samples and positive controls containing dye blobs within the defined analysis range (60– 460bp) do not need to be re-injected or re-run if the dye blob does not interfere with the sizing of the sample or data interpretation.

6.2.1.3.4. Pull-up (Bleed-through) – pull-up peaks are caused by one peak “bleeding” into another color. Unusually high pull-up (>3%) or extra peaks can be caused by oversaturation of the system or a problem with the spectral. Pull-up peaks are a result of the spectral not correcting for all the spectral overlap. These peaks are easily

recognized by overlaying the colors and observing the alignment of peaks at (typically +/- 0.25bp) the same data point or evaluating the sample's raw data. Samples and controls exhibiting pull-up peaks do not need to be re-injected or re-run if it does not interfere with the sizing of the samples. When an ongoing spectral problem exists, a new spectral must be run, applied and the samples re-injected/re-run.

6.2.1.3.5. Split Peaks/Minus A/Non-template Nucleotide Addition – split or minus A peaks may be seen in a sample if amplification did not go to completion. Taq polymerase catalyzes the addition of a single nucleotide (predominantly adenosine) to the 3' ends of double stranded PCR products. This non-template addition results in a PCR product that is one base pair longer than the actual target sequence (+A). Split peaks and minus A are a result of the incomplete addition of an adenosine, where each allele is represented by two peaks one base pair apart (-A and +A). Split or minus A peaks may be observed when the amount of input DNA is greater than recommended. This occurs because more time is needed for the Taq polymerase to add the A nucleotide to all molecules with each cycle. Samples with apparent excess PCR product may be re-amplified using less input DNA.

6.2.1.3.6. Shoulders – shoulder peaks are approximately 1 – 4bp smaller, but occasionally larger than the main allele, caused by a flat decline of the fluorescence of the main peak. Shoulders are easily recognized because they do not have proper peak morphology. Samples and controls exhibiting shoulders do not need to be re-injected or re-run if it does not interfere with the sizing of the samples.

6.2.1.3.7. Elevated Baseline- raised baseline will appear as non-specific elevation of the baseline causing signal(s) to be labeled. These labeled signals do not typically display proper peak morphology. Samples and controls exhibiting elevated baseline do not need to be re-injected or re-run if it does not interfere with the sizing of the samples.

6.2.1.3.8. GlobalFiler™ Kit Artifacts- artifacts and anomalies have been seen in data produced using the

GlobalFiler™ Amplification Kits, in general, and during QC of specific kit lot numbers. Artifacts are typically reproducible while anomalies (such as spikes or baseline noise) are not observed consistently and are typically not reproducible. Analysts should consider these artifacts and anomalies when interpreting data. In general, samples and controls exhibiting these types of non-specific artifacts and/or anomalies do not need to be re-injected or re-run. Refer to the GlobalFiler™ PCR Amplification Kit User Guide revision E, Chapter 5; the Artifacts Identified Post-Developmental Validation: GlobalFiler™ PCR Amplification Kit Technical Note (current version); and the DFS Internal Validation of the GlobalFiler™ PCR Amplification Kit for examples of kit specific artifacts

6.2.1.3.9. Off-scale Peaks/Data – an off-scale peak occurs when the fluorescent signal exceeds the dynamic detection range of the CCD camera in the 3500/3500xL Genetic Analyzer. Capillary electrophoresis instruments are limited in the amount of fluorescence they can detect. Signal saturation will be indicated by the presence of a pink line through the allele(s) and the corresponding GMID-X quality flag. The GMID-X software is unable to determine the true peak heights for these alleles. Their values should not be used for quantitative evaluation (e.g., heterozygote balance calculations or stutter calculations). In addition, off-scale data may cause many of the above listed artifacts. As needed, samples containing off-scale data should be re-amplified at a lower target concentration.

6.2.1.3.9.1. Some samples may appear to be overloaded without displaying a GMID-X quality flag. At the discretion of the analyst, these may be either edited or the sample re-amplified at a lower concentration.

6.2.1.3.9.2. Evidence mixture profiles that contain an STR locus with peak(s) above 25,000 rfu will be reamplified and rerun using a decreased template amount.

6.2.2. **Overall Profile Quality** – the overall quality of each profile, as well as the potential number of contributors and any relevant source information, will be assessed prior to any in-depth determinations about the data observed. The presence of apparent degradation, inhibition, or significant stochastic effects in a profile will be considered during the interpretation process and may influence the assessments made at each locus. In addition, the number of potential contributors to a sample will affect how the sample is interpreted. It is also possible that knowledge or perceived knowledge regarding a potential contributor(s) to a mixture sample may affect sample interpretation (e.g., an intimate swab with semen positive results). These observations will be documented in the case file (e.g., electropherogram and/or STRmix report).

6.2.2.1. Degradation: DNA degradation is a process in which DNA molecules randomly break down into smaller pieces. This becomes problematic for forensic DNA typing when it occurs within the target DNA sequences of PCR. If the DNA has been degraded or damaged at these locations, sister peak height imbalance, abnormally high stutter product(s), allele dropout, and/or locus dropout may be observed. Typically, degraded profiles will display a downward sloping or "ski-slope" pattern from left to right for each dye. This pattern is due to the larger/longer loci being more likely to contain a break due to degradation than the smaller loci. It is important to note that it is possible for the profile from one or more donors in a multiple contributor profile to exhibit degradation in the absence of observed profile degradation of the other donor(s).

6.2.2.2. Inhibition: DNA samples may contain one or more of the PCR inhibitors commonly encountered in forensic casework. The presence of inhibitors may manifest itself by the failure to produce results at some or all loci and may mimic results seen when degradation is present. Samples containing inhibitors often produce partial profile results in which the smaller loci drop out before the larger loci.

6.2.2.3. Within STRmix™ v2.4, dropout is modeled as an extreme form of imbalance.

6.3. Determination of Single Source or Mixture Profile

6.3.1. General Single Source Profile Attributes: A single source profile will typically have no more than two relatively well-balanced alleles (i.e. Heterozygote balance (Hb), $0.55 \leq Hb \leq 1.82$) at all loci. In rare instances, an individual may exhibit a tri-allelic pattern at a locus. When a tri-allelic pattern is noted and not confirmed by multiple observations in a case, the sample will typically be re-amplified to confirm its presence.

6.3.1.1. When a tri-allelic pattern is noted, the results will be denoted, though the locus will be ignored in the STRmix™ analysis.

6.3.2. Following determination that a profile is single source, all of the profile data, as well as the data at each locus, is assessed to determine whether the result(s) may be used for comparisons. Single source profiles exhibiting results at less than 5 autosomal loci and average peak heights of approximately 125 rfu or less will be deemed uninterpretable and reported as such.

6.3.3. General Mixture Profile Attributes: A mixture profile typically has more than two alleles at one or more loci (with a tri-allelic pattern as the exception) and/or exhibits significant imbalance between peak heights at one or more loci. The presence of more than two peaks at one locus or imbalanced peak heights does not, in and of itself, necessitate a profile being designated as a mixture. While it may be possible to infer a sample is a mixture based on > 2 alleles/peaks and/or significant peak height imbalance at a single locus, it is preferable to observe characteristics consistent with a mixture at no fewer than two loci.

NOTE: As a guideline, peak height imbalance would be assessed using the heterozygote balance (Hb) formula as listed in the Calculations section 8.1 ($0.55 \leq Hb \leq 1.82$).

6.3.4. Following determination that a profile consists of a mixture, all of the profile data, as well as the data at each locus, is assessed to determine whether the result(s) may be used for comparison.

6.4. Determination of the Number of Contributors

6.4.1. Generally, the minimum number of contributors to a mixed sample can be determined based on the locus that exhibits the greatest number of allelic peaks. For example, if at most five alleles are observed at a locus, then the DNA results are consistent with having arisen from at least three individuals. Proceed with caution when only one allele in the entire sample would lead to an increased number of possible contributors, as stutter, imbalance, and the potential for a tri-allele can complicate mixture interpretation. Loci more likely to exhibit a tri-allele pattern are D18S51, D21S11, FGA, TPOX, and vWA. While counting allele peaks is very useful in determining a minimum number of contributors, the analyst must also consider that allele sharing between individuals may result in an underestimate of the actual number of contributors.

6.4.2. Minimum Number of Contributors Assessment: The minimum number of contributors to a mixture will generally be determined by examining the locus demonstrating the largest number of peaks ≥ the analysis method AT parameters. The peaks are counted at that locus and then divided by two. The resulting quotient is rounded up to the nearest integer for the minimum number of contributors.

6.4.2.1. In general, a profile containing a mixture of DNA from two people will have no more than four alleles at all loci; a profile containing a mixture of DNA from three people will have no more than six alleles at all loci; and so on (i.e., there will be no more than 2 x [the number of contributors] alleles at all loci).

6.4.2.2. Ultimately, the number of alleles and their rfus will be used to more accurately assess the number of potential contributors to a mixed profile.

6.4.2.3. Low-level profiles (especially at loci with alleles that exhibit stochastic effects) may not exhibit more than two alleles at any locus but information gleaned from the profile as a whole may allow for the profile to be characterized as a mixed sample.

6.4.2.4. Conversely, significant peak height imbalance at multiple loci in samples amplified with the target amount of input DNA may provide information regarding the potential for an additional contributor(s) to a mixture profile (e.g., as in mixtures containing DNA from two or more biologically related individuals).

6.4.2.4.1. Peak height imbalance can be assessed using heterozygote balance (Hb). Refer to the calculation listed in section 8.1. In some cases, genotype combinations may be paired based on heterozygote balance to further determine the number of contributors.

6.4.2.4.2. It is noted that it may not be possible to accurately assess the potential number of contributors to mixture profiles exhibiting characteristics consistent with having at least 3 contributors.

6.4.2.5. Indistinguishable Number of Contributors: Sometimes the number of contributors may be unclear due to the complexity of the profile. If an analyst is unclear as to the number of contributors present in a GlobalFiler™ mixture, and the data supports both assessments, the analyst will run STRmix™ analysis as follows:

6.4.2.5.1. 2 or 3 contributors: Run the mixture as both 2 and 3 person contributor profiles.

6.4.2.5.2. 3 or 4 contributors: Run the mixture as both 3 and 4 person contributor profiles.

6.4.2.5.3. 4 or 5 contributors: Run the mixture as both 4 and 5 person contributor profiles.

6.4.2.6. Uninterpretable Profiles: Based on limited data and/or complexity of a mixture, the following types of profiles will be deemed uninterpretable in their entirety:

6.4.2.6.1. 5 or 6 contributors: The profile will be deemed uninterpretable.

6.4.2.6.2. 6 or more contributors: The profile will be deemed uninterpretable.

6.4.2.6.3. The profile is low level and the number of contributors cannot be assessed: The profile will be deemed uninterpretable.

6.5. Application of Assumptions for Known Contributors

6.5.1. Assumptions are valid when a known contributor is presumed to be present on an intimate item. Intimate samples are defined as biological evidence samples taken directly from an individual's body (e.g., vaginal swabs, swabs from a bite mark on skin, fingernail scrapings).

6.5.2. Assumptions are also valid when a known contributor is presumed to be present on a personal item(s) (e.g., underwear, clothing, bedding, steering wheel of personal vehicle) as long as there is reasonable expectation of ownership/usage.

6.5.3. The presence of an additional known contributor (e.g., consensual partner) may also be made when deemed appropriate.

6.5.4. All assumptions must be clearly stated in the Report of Examination.

6.5.5. The analyst will make a visual (qualitative) comparison between the assumed contributor profile and the evidence profile to ensure the assumed contributor is represented in the evidence profile.

6.5.6. With respect to intimate samples, failure to detect the full genotype of the assumed contributor at all loci or at a locus will not be automatic cause for exclusion. Indeed, for intimate samples potential assumed contributor genotype information/results will be critically reviewed to ensure interpretations 'make sense'. In general, detection of an individual's genotype on a sample obtained from their body is considered a non-probative result.

6.6. When to Perform STRmix™ Analysis

6.6.1. Single Source Profiles:

6.6.1.1. All probative single source profiles that match a known reference will be uploaded for STRmix™ analysis if results are obtained at 5 autosomal loci or more and average peak heights are generally greater than 125 rfu. Refer to FBS32 – GlobalFiler™ Data Analysis Using STRmix™.

6.6.2. Two Person Mixture Profiles:

6.6.2.1. All 2 person mixture profiles need to be deconvoluted using the STRmix™ software. Refer to FBS32 - GlobalFiler™ Data Analysis Using STRmix™. The following are exceptions:

6.6.2.1.1. The profile has been deemed uninterpretable (Section 6.4.2.6.).

6.6.2.1.2. Based on CODIS eligibility, within the context of available case information, non-CODIS eligible mixture profiles without an associated probative reference typically will not be deconvolved.

6.6.2.1.2.1. If an inclusionary reference sample is provided at a later date, a deconvolution will then be performed.

6.6.2.1.2.2. The upper level contributor is non-probative AND the lower level contributor displays less than ten unique minor alleles at the autosomal loci and average peak heights of less than 125 rfu.

6.6.3. Three, Four, and Five Person Mixture Profiles:

6.6.3.1. All 3, 4, and 5 person mixture profiles need to be deconvoluted using the STRmix™ software. Refer to FBS32 – GlobalFiler™ Data Analysis Using STRmix™. The following are exceptions:

6.6.3.1.1. The profile has been deemed uninterpretable (Section 6.4.2.6).

6.6.3.1.2. Based on CODIS eligibility, within the context of available case information, non-CODIS eligible mixture profiles without an associated probative reference will not be deconvoluted.

6.6.3.1.3. If an inclusionary reference sample is provided at a later date, a deconvolution will then be performed.

6.6.3.2. All 4 and 5 person mixture profiles require the use of replicate amplifications (same target template) for STRmix™ analysis.

6.6.3.2.1. Based on case scenario or other profile results, an analyst may choose not to replicate a 4 or 5 person mixture; however, this profile will be deemed uninterpretable.

6.6.4. Single source, 2 and 3 person mixture profiles may use replicates at the discretion of the analyst.

6.7. Evaluating the STRmix deconvolution

6.7.1. Following the evaluation the STRmix™ data of in FBS32, an analyst may determine that a 3, 4, or 5 person mixture profile or a specific contributor of that mixture is not suitable for comparison based on the following guidelines determined by validation data:

6.7.1.1. For three and four person mixtures, contributors with DNA amounts less than 200 may produce false inclusions.

6.7.1.2. For five person mixtures, contributors with DNA amounts less than 375 may produce false inclusions.

6.7.1.3. If the mixture sample has been replicated, it is possible that the DNA amounts listed above are lower.

6.7.1.4. While the values above are listed as specific numbers, they represent approximate levels at which the data may be limited.

6.7.2. Analysts may choose to interpret a contributor or profile despite the DNA amount assigned depending on certain circumstances of the data, sample type, or case information.

6.8. Comparison of Interpreted Profiles to Probative Reference Samples

6.8.1. After interpretation (including STRmix™ deconvolution, as applicable) has been completed on all case profiles, qualitative comparisons to probative reference samples submitted in connection with the investigation are conducted, if applicable.

6.8.1.1. Inclusions: Comparison of data (electropherogram and STRmix™ deconvolution, if applicable) between an unknown sample and a known sample are consistent with each other. A known sample may be included to a single source sample full (complete profile) or partial (results not obtained at all loci), or included into a mixed sample.

6.8.1.2. Exclusions: Comparison of data (electropherogram and STRmix™ deconvolution, if applicable) between a known and unknown sample are not consistent with each other.

6.8.1.3. Inconclusive: There is not enough data to support an inclusion or exclusion.

6.8.2. When an individual is qualitatively included as the source of a single source profile and the result is probative, a statistical calculation is conducted using the STRmix™ software.

6.8.3. When an individual is qualitatively included as a contributor to a mixture profile and the result is probative, or if the qualitative comparison is inconclusive and the result is probative, a statistical calculation is conducted using the STRmix™ software.

6.8.4. When an individual is qualitatively excluded as a contributor to a DNA profile/mixture, a statistical calculation will not be performed.

NOTE: A pseudo-exemplar or alternate known that exhibits a mixture will not be used for comparisons.

6.9. Statistical Analysis

6.9.1. Statistical calculations will be performed using the STRmix™ software, which uses a fully continuous probabilistic genotyping likelihood ratio (LR) model.

6.9.2. The LR is calculated using the allele frequencies referenced the 2015 Expanded FBI STR Population data set (Caucasian, African American (combined), Southwest Hispanic and Southeast Hispanic) and the Balding and Nichols formulae (NRC II Recommendation 4.2 in "The Evaluation of Forensic DNA Evidence" (1996)) provided in Appendix 13 of FBS30 – GlobalFiler™ Data Analysis Using STRmix™. The theta value of 0.01 is used in the Balding and Nichols formulae.

6.9.2.1. The allele frequencies are then adjusted in STRmix™ using a Bayesian posterior mean frequency to better account for sampling uncertainty with allele counts within a limited population data set. This formula is listed in the Calculations section 8.2.

6.9.2.2. When an allele has not been observed in the database, the posterior mean frequency (f_i) reduces to the formula listed in the Calculations section 8.3.

6.9.3. The combined LR calculated by STRmix™ is referred to as a point estimate. Because the true answer is not known, a credible interval is calculated using the Highest Posterior Density (HPD) method and then applied around the point estimate. This interval accounts for the uncertainty associated with the point estimate LR. The lower bound of the HPD interval is reported from STRmix™ to be conservative to the accused or person of interest.

6.9.4. Statistical calculations will only be reported on evidentiary material, when deemed to be relevant in the context of the case.

6.9.5. If a sample is run twice in STRmix™ using two different numbers of contributors, report the lowest likelihood ratio (LR) as "at least" from the four population groups used. The number of contributors used to calculate the LR will be documented in the report.

6.9.6. Likelihood Ratio (LR)

6.9.6.1. The LR is represented by the following equation:

$$LR = \frac{Pr(E \mid Hp)}{Pr(E \mid Hd)}$$

6.9.6.2. A LR requires the evaluation of the probability of the evidence under two competing propositions or hypotheses.

6.9.6.2.1. Hp = the Hypothesis of the prosecution. A sensible proposition will be chosen based on available case information.

Hd = the Hypothesis of the defense. A sensible proposition will be chosen based on available case information and will typically be the opposite of the Hp.

Pr = Probability

E = Evidence

| = Given

6.9.6.3. A point estimate for the LR with associated weighted genotype combinations is then calculated when known contributors are compared to the evidence. The general formula used is listed in the Calculations section 8.4.

6.9.7. Formulating Propositions for Calculating the Likelihood Ratios

6.9.7.1. Multiple pairs of propositions (other than different number of contributors) may be required in certain instances. This information will be documented in the case file. If multiple runs are performed in STRmix™ for multiple pairs of propositions that are evaluated, maintain all analyses in the case file. The analyst will determine which pair of proposition(s) to report based on available case information.

6.9.7.2. Propositions are mutually exclusive (both cannot be true at the same time).

6.9.7.3. The number of contributors in both the Prosecution and the Defense hypotheses will be the same based on the analyst's interpretation obtained from the DNA data.

6.9.7.4. Single Source Profiles:

6.9.7.4.1. Set the Hp (Prosecution hypothesis) first. For single source profiles, the hypothesis may typically include the Person of interest (POI), Victim (V), or other known single individual or unknown individual (U).

6.9.7.4.2. Set the Hd (Defense hypothesis) second. For single source profiles, the hypothesis may typically be unknown (U).

6.9.7.4.3. An example of a pair of propositions for a single source sample is POI/U.

6.9.7.5. Mixture Profiles:

6.9.7.5.1. Set the Hp (Prosecution hypothesis) first. List all contributors that are possible in the numerator. Possibilities may include the Victim (V), the person of interest (POI), any consensual partners (C), any

additional knowns (e.g., witnesses) or any additional unknowns (U).

6.9.7.5.2. Set the Hd (Defense hypothesis) second. List all contributors that are possible in the denominator. Possibilities may include the Victim (V), the person of interest (POI), any consensual partners (C), any additional knowns (e.g., witnesses) or any additional unknowns (U).

6.9.7.6. Conditioning: Incorporates an assumption that a person can reasonably be included in the profile. Example: the victim can be assumed on her own vaginal swab.

6.9.7.6.1. The conditioned individual will appear in both the numerator and the denominator. More unknowns in the denominator compared to the numerator generally increase the LR.

6.9.7.6.2. Examples of Conditioning with Mixture Samples:

6.9.7.6.2.1. Two person mixture profile explained by Victim and Person of Interest on a vaginal swab. The possible pairs of propositions include:

1) V+POI/V+U

or

2) V+POI/U+U

The first pair of propositions is more reasonable, as one can expect the victim to be on her own vaginal swab.

6.9.7.6.2.2. Three person mixture profile explained by Victim, Consensual partner and a Person of Interest on a vaginal swab. The possible pairs of propositions include:

1) V+C+POI/V+C+U

or

2) V+C+POI/V+U+U

or

3) V+C+POI/U+U+U

The first pair of propositions is most reasonable, as one can expect the victim and consensual partner on the victim's vaginal swab.

6.9.7.7. Neutral Evidence: Evidence that cannot be attributed to an individual, such as an item found in a public area (e.g. park, street). In these circumstances, examine all possible pairs of propositions.

6.9.7.7.1. Choose the pair of proposition(s) most reasonable based on the available case information. Be prepared that the pair of propositions could change.

6.9.7.7.2. No conditioning will be conducted on neutral evidence.

6.9.7.8. Multiple pairs of propositions may be run on STRmix™. If multiple pairs of propositions are run, all analyses will be maintained in the casefile.

7. Sampling

7.1. Not applicable

8. Calculations

8.1. Heterozygote Balance (Hb) formula:

$$Hb = \frac{O_{HMW}}{O_{LMW}}$$

O = observed peak height

HMW = high molecular weight

LMW = low molecular weight

8.2. Bayesian posterior mean frequency formulae:

$$f_i = \frac{(\text{observed allele frequency})(N) + 1/k}{N+1}$$

where f_i = posterior mean frequency

N = total number of alleles observed at the locus

k = number of possible alleles

8.3. Posterior mean frequency for alleles not observed in the sampling database:

$$\frac{1/k}{N+1}$$

8.4. A point estimate for the likelihood ratio (LR) with associated weighted genotype combinations formula:

$$LR_C = \frac{\sum_j w_j \Pr(S_j \mid H_p)}{\sum_u w_u \Pr(S_u \mid H_d)}$$

8.4.1. $\sum_j w_{j\ Pr(Sj|Hp)}$ = sum of the probabilities of all possible genotypes (S_j) given the prosecution's hypothesis (Hp) (i.e., where the POI is included) with weightings w_j applied.

8.4.2. $\sum_u w_{u\ Pr(Su|Hd)}$= sum of the probabilities of all possible genotypes (S_u) given the defense's hypothesis (Hd) with weightings w_u applied for an unknown contributor, unrelated to the POI.

9. Uncertainty of Measurement

9.1. Sampling uncertainty occurs due to the finite allele probabilities associated with the population samples being used. In STRmix™ an allowance for sampling uncertainty is implemented by adjusting the allele frequencies using a Bayesian posterior mean frequency (i.e., Highest posterior density (HPD)) to better account for sampling uncertainty with allele counts within a limited population data set.

9.2. The combined LR calculated by STRmix™ is referred to as a point estimate. Because the true answer is not known, a credible interval is calculated using the Highest Posterior Density (HPD) method and then applied around the point estimate. This interval accounts for the uncertainty associated with the point estimate LR. The lower bound of the HPD interval is reported from STRmix™ to be conservative to the accused or person of interest.

9.3. STRmix™ uses the Balding and Nichols model (NRC II recommendation 4.2) which applies a co-ancestry coefficient (Θ), a measure of the probability that two alleles taken from two individuals of the same sub-population are identical by descent.

10. Limitations

10.1. Not applicable

11. Documentation

11.1. The following documents may be generated during the interpretation process:

11.1.1. STRmix™ Advanced Report

11.1.2. FBU Report of Examination

12. References

12.1. STRmix™. Version 2.4 User's Manual (Current Version)

12.2. DFS STRmix™ v2.4 Internal Validation Report: Part 1 and 2 (2016-2017).

12.3. DFS The Zoom Study: Additional Guidelines for Interpretation of Mixtures and Low Level Data Using GlobalFiler™ on the 3500/3500xL and/or STRmix™ 2.4 (08/10/2020).

12.4. Balding, D.J. and R.A Nichols, DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands. Forensic Science International, (1994). 64: 125-140.

12.5. National Research Council. The Evaluation of Forensic DNA Evidence, Washington, DC: Academy Press, 1996. (colloquially referred to as "NRC II").

12.6. Applied Biosystems. GlobalFiler™ PCR Amplification Kit User's Manual Applied Guide (current revision).

12.7. Quality Assurance Standards for Forensic DNA Testing Laboratories, Federal Bureau of Investigation (current revision).

12.8. SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories (current revision).

12.9. ISO/IEC 17025 General Requirements for the Competence of Testing and Calibration Laboratories, International Organization for Standardization, Geneva, Switzerland (current revision)

12.10. ANSI-ASQ National Accreditation Board Supplemental Requirements for Forensic Test Agencies, Milwaukee, WI (current revision)

12.11. GlobalFiler™ Data Analysis Using GeneMapper ID-X 1.6 (FBS39)

12.12. GlobalFiler™ Data Analysis Using STRmix™ 2.4 (FBS32)

12.13. FBU Report Wording (FBS26)

12.14. Artifacts Identified Post-Developmental Validation: GlobalFiler PCR Amplification Kit Technical Note (current version)